

修 士 論 文 の 和 文 要 旨

研究科・専攻	大学院 情報理工研究科 情報学専攻 博士前期課程		
氏 名	楊 暁旭	学籍番号	1830117
論 文 題 目	雑談対話における弱関連話題への転換手法		
要 旨			
<p>近年、特定の要求への情報提供を目的とせずに、主に雑談形式の対話を目指す非タスク指向型雑談対話システムが注目されている。その中で、雑談を行う際に延々と同じ話題を続けていると、ユーザは対話システムとの雑談に飽きてしまう可能性が高い。これを回避するためには、雑談対話システムに話題転換の機能を持たせる必要がある。そこで本研究では、雑談対話における話題転換を対象として、発話文解析、話題転換点の推定、転換話題の決定、発話文の生成という4つの処理を考える。その中で、話題転換点の推定と転換話題の決定という2つの処理手法を新たに提案する。話題転換点の推定では、ユーザとシステムの発話文ペアを用いる先行研究の方法と異なり、ユーザの発話文間の類似度を計算することで話題転換点の判断を行う。転換話題の決定では、2つの手法を提案する。1つ目の方法では、発話文中の話題語に基づいて話題候補語集まる。話題語候補の各単語と発話文中の全て名詞との類似度を計算し、平均類似度が最も小さい話題語候補を弱関連話題と決定する。もう1つの方法では、発話文の話題語を含 tri-gram と同じ構造を持つ話題語候補の出現頻度によって、弱関連話題を推定する。提案した手法を実装し、実際の雑談会話データに対して、2種類の評価実験を行う。評価実験1では、本手法から推定した話題転換点と人手で評価した話題転換点を比較して、具体的な結果（適合率と再現率の平均値）を示すことである。評価実験2では、雑談対話中で、転換が必要な場合で弱関連話題がランダム話題により、興味度、適切さ、話題語含む転換文の自然さをそれぞれ五段階評価する。その結果は先行研究より本手法が悪かったと分かった。今後の課題としては、ユーザの発話文の発話行為の分類による話題転換点の精度の向上、弱関連話題の多様性の工夫による転換話題の興味度と適切さの向上が上げられる。</p>			

令和元年度 情報学専攻 修士論文

雑談対話における弱関連話題への転換手法

提出年月日： 令和 2 年 1 月 31 日

提出者： 学籍番号 1830117

楊 暁旭

指導教員： 内海 彰 教授

松吉 俊 助教

目次

1	はじめに	3
2	関連研究	5
2.1	話題転換の分類	6
2.1.1	雑談対話システムにおける話題展開	6
2.2	雑談対話システムにおける話題転換	7
3	提案手法	9
3.1	発話文解析	9
3.1.1	形態素の抽出	10
3.1.2	有効発話文の選択	10
3.2	話題転換点の推定	10
3.2.1	話題転換点の推定	11
3.2.2	話題転換類型の推定	12
3.3	転換話題の決定	12
3.3.1	話題候補語の平均類似度による手法	13
3.3.2	tri-gram の構造を用いた手法	13
3.4	話題転換文の生成	14
4	評価実験	16
4.1	実験方法	16
4.1.1	話題転換点の評価実験 (実験 1)	17
4.1.2	転換話題の評価実験 (実験 2)	19
4.2	実験結果	19
5	考察	22
5.1	話題転換点の評価	22
5.1.1	話題展開点の分類	22
5.1.2	話題展開点の考察	28

5.2	転換話題の考察	29
5.2.1	話題展開点の類型によって転換話題の考察	31
5.2.2	不適切な話題語の排除	32
5.3	今後の課題	34
6	おわりに	35
	参考文献・謝辞	36
	付録 A 付録	39
A.1	ストップワード	39
A.2	Twitter の話題	39
A.3	話題転換点の個数と対話履歴範囲	40
A.4	類似度平均値で決定した弱関連話題	41
A.5	tri-gram で決定した弱関連話題	45

1 はじめに

近年，人間と対話をするコンピュータのことを対話システムという．対話システムは大きく二種類に分けられる．一つはタスク指向型対話システム，もう1つは非タスク指向型対話システム（雑談対話システム）である．前者は対話を通して特定のタスクを遂行するシステムのことで，Siri^{*1}はその典型例である．後者はよく言われるチャットボットであり，例えば，りんな^{*2}などの対話の目的が明らかなでないシステムである．その中で，何らかの目的を持って対話を行うタスク指向型対話システムに関する研究は顕著な成果を収めている [1]．しかし，特定の要求への情報提供を目的とせず，主に雑談形式の対話を目指す非タスク指向型雑談対話システムがまだ発展途上の段階にあると言える．その原因の1つとして，タスク指向型においては限られた範囲での話題を扱えばよいのに対して，非タスク指向型においては多様な話題を扱う必要があるため，個々の話題に対する準備が不十分になってしまうことが挙げられる [2]．雑談対話システムとの対話を長く継続させるためには，話題の内容を深めたり，新たな話題を提供することが必要となる．雑談を行う際に延々と同じ話題を続けていると，ユーザは対話システムとの雑談に飽きてしまう可能性が高い．これを回避するためには，雑談対話システムに話題転換の機能を持たせることが必要である．

雑談対話システムにおける話題転換には話題転換点の推定，転換話題の決定，発話文の生成という3つの処理が必要である [3]．本研究では，話題転換点の推定と話題転換の決定を研究対象として，新たな手法を提案する．その理由は，発話文の生成より話題転換点の推定や転換話題の選択が重要だからである．不適切な話題転換点で，または不適切な転換話題で転換すると，ユーザに唐突感や違和感を感じさせ，対話への興味が失われる可能性が高い．

藤本ら [4] は転換前後の話題の関連度によって，関連性の強い話題への転換（強関連話題），中程度の関連性をもった話題への転換（中関連話題），関連性がほとんどない話題への転換（弱関連話題）の3つのタイプに分類している．話題転換に関するほとんどの先行研究 [5, 6, 7] は，強関連話題または中関連話題の話題転換に関する研究である．

しかし，弱関連話題の転換話題が人間同士の間でもよく起こる [4] ことから，ユーザが対話システムとの雑談を継続するために重要であると考えられる．特に，ユーザが先行話題に飽きている時に，弱関連話題に転換することで，ユーザの対話への興味を喚起することが期待できる．弱関

*1 <https://www.apple.com/jp/siri/>

*2 <https://www.rinna.jp>

連話題の話題転換に関する研究としては、堀内ら [8]、前田ら [9] と瀧口ら [10] がある。しかし、堀内ら [8] と前田ら [9] の手法では、システム発話とユーザ発話の一致度、単語数、連続性という指標で話題転換点を推定し、転換話題をウェブからランダムに決定するだけであり、不自然に感じる話題転換になってしまう可能性がある。瀧口ら [10] は対話が破綻しているかを検出して、破綻している場合に話題転換点を見なす。そして、検索エンジンの中で検索対象に対して、閲覧数などの時間と共に変化する動向情報によって転換話題を推定する。また、特徴的な変動のタイプとして、最大値、最小値、ピーク、ボトム、急上昇、急降下の 6 種類のうちいずれかを指定して検索できる。具体的には、同様の特徴的変動が見られた時期に、直前までの話題と関連アイテム候補のどちらもキーワードとして含むツイートが最も多い特徴量を関連性がある単語を転換話題を推定すると、固有名詞が転換話題となることがあり、不自然を感じさせてしまうことがある。

そこで本研究では、ユーザの連続発話を考慮し、適切な話題転換点を推定する。また、発話文の名詞に基づいて弱関連話題を推定して、ユーザの対話興味を喚起できる話題転換の実現を目的とする。

2 関連研究

話題転換の定義では，藤本ら [4] は会話の流れに無関係な話題に遷移することは少ないという考えから，話題・発話間のつながりの自然さに着目し，話題転換の自然さを Type A, Type B, Type C の三段階に分類している．そして，一つの話題ブロック（同一の話題が継続されている発話区間を話題ブロックと呼ぶ）が一定ターン以上継続した場合を対話が活性化された状態であるとして，その話題ブロックの起点となった話題転換 Type A, Type B, Type C の個数と式 (2.1) による平均活性化率の計算結果表 2.1 に示す．

$$\text{活性化率} = \frac{\text{Type } X \text{ への活性話題転換点数}}{\text{Type } X \text{ への話題転換点の総数}} \quad (2.1)$$

以上によって，雑談対話中で，弱関連話題の話題転換が個数と活性化話題の役割ともにが相当であると分かった．本研究では，藤本ら [4] の話題分類によって，強関連話題（Type A），中関連話題（Type B），弱関連話題（Type C）という 3 つ類型に転換話題を分けている．その中で，弱関連話題の話題転換の決定方法を提案する．また，藤本ら [4] も話題転換点で話題転換文と先行話題の関連性を探求するために EDR 電子化辞書を用いて単語と発話間の概念的関連性（以下，拡張概念と呼ぶ）を計算している．分析の結果，話題転換点での発話と話題のベクトルの類似度には統計的な傾向がみられた．しかしながら話題転換点および転換タイプを発話が行われた時点で判別する指標としては不十分であった．これは単語と概念表現の対応の曖昧性や一発話から得られる情報の少なさ，および拡張概念であっても人間の持つ概念的関連性の知識を完全には表現できないために省略されている語を介しての関連性を適切に捉えられていないことが原因と考えられる．

表 2.1 転換話題の種類ごとの有効性

転換話題類型	活性話題個数	平均活性化率
Type A	11	0.47
Type B	156	0.69
Type C	47	0.50

2.1 話題転換の分類

藤本ら [4] は転換前後の話題の関連度によって、関連性の強い話題への転換 (TypeA), 中程度の関連性をもった話題への転換 (TypeB), 関連性がほとんどない話題への転換 (TypeC) の3つのタイプに分類している。瀧口ら [10] もこの分類に基づいて話題転換を分類している。本研究では、話題展開 (Type A, Type B) と話題転換 (Type C) の2つのタイプに分類する。

2.1.1 雑談対話システムにおける話題展開

転換前話題と転換後話題の間にある程度の関連性を持つ場合を話題展開と呼ぶ。つまり、関連性の強い話題への転換 (TypeA) と中程度の関連性をもった話題への転換 (TypeB) である。話題展開に関する先行研究として、文献 [5, 6, 7] がある。中野ら [5] は発話文中の名詞のみを話題語として着目し、単語分散表現の学習に word2vecf を用いて、短時間で効率的に大規模データに対して学習させる。発話生成の際に名詞に対して単語分散表現を用いて関連した単語に入れ替えるという話題展開を提案している。浅見ら [6] は大規模コーパスから導出された単語分散表現に基づいて連想による話題展開手法を提案している。また三上ら [7] は対話システムにおける話題の展開手法に対して、実数値遺伝的アルゴリズムと複数の評価関数を組み合わせて分散表現の変換を行うことで、あいまい性を考慮し未知語や未学習語に対して高い柔軟性を保ちつつ、より多様な方向へ話題を導けるような話題語の変換を行うシステムを提案した。以下は話題展開の対話例である。

話題展開例：

...

System : ご飯はどうですか？

User : 大学の食堂で食べる。

System : へー。

User : 銭湯も同じです。

System : あーあー、そうですか。

User : うん。

System : なるほど、ところで、温泉が好きですか？

User : 大好きです。

2.2 雑談対話システムにおける話題転換

話題転換では、それまで話題となっていた対象や事態とは異なる、新しい対象や事態への言及になる。つまり、関連性がほとんどない話題への転換（TypeC）である。Li ら [3] の研究に従って、雑談対話システムにおける話題転換を話題転換点、転換話題の決定、転換文の生成という 3 種類の処理で構成する。話題転換を行った先行研究として堀内ら [8] や前田ら [9] が挙げられる。これらの研究では、話題転換点の確定とリアルタイムでの話題生成という 2 つの処理方法を提案している。話題転換点の推定手法では、堀内ら [8] はシステム発話文に対するユーザ発話文の内容語数の個数、ユーザ発話文とシステム発話文の一致度、システムから連続する発話文に対するユーザの反応等の指標から閾値を導き、転換話題点を確定する。前田ら [9] は堀内ら [8] の手法に基づいて、話題展開について改良しただけであり、話題転換のタイミングを決める転換指標の考慮がまだ不十分であると考えられる。また、転換話題の生成では、堀内ら [8] と前田ら [9] は新しい話題を kizasi.jp^{*1} から取得し、ユーザに提供する。しかし、10 分ごとに話題のランキングが変化していくため、生成した話題がランダム性を持っていると見なすことができる。そのため、この手法で生成した話題を弱関連話題として話題転換をすると、ユーザに唐突感や違和感を感じさせ、対話への興味を失わせる可能性があると考えられる。

瀧口ら [10] の研究では、話題転換点の推定に対話破綻の検出の技術を用いている。破綻した応答をエージェントが返した際はユーザが応答を返すまでの時間が長く、また応答そのものも短くなる傾向があるため、それらを話題転換点の指標と見なしている。また話題転換については、中関連話題に転換することが望ましいと仮定する。この考えに基づき、動向情報（例えば、Wikipedia の閲覧数のように時間とともに変化するデータ）とソーシャルメディア（Twitter 情報）のデータからある程度の関連性を持つ話題を発見する手法を提案している。しかし、瀧口ら [10] の研究で提案された文字当たり応答時間については、自然な発話と不自然な発話の差が有意であることを示したが、破綻していない通常の対話時における、話題転換を行うための適切なタイミングを判断する指標がまだない。さらに、対話破綻時の話題転換による対応については話題転換の対象となる話題が見つからない状況がある。また、固有名詞を話題とした話題転換を行った場合対話が破綻してしまうこともある。

そこで本研究では、雑談対話システムにおける話題転換の中で、ユーザの連続発話文を考慮し、適切な話題転換点を推定する。また、発話文の話題語に基づいて弱関連話題を推定して、ユーザ

^{*1} <https://kizasi.jp/>

の対話興味を喚起できる話題転換の実現を目的とする．さらに本研究では，日常対話データを利用して，提案手法の有効性を評価させる．以下は話題転換の対話例である．

話題転換例：

...

System：ご飯はどうですか？

User：大学の食堂で食べる。

System：へー。

User：銭湯も同じです。

System：あーあー、そうですか。

User：うん。

System：なるほど、ところで、専攻は何ですか？

User：日本語です。

3 提案手法

本研究では，Li ら [3] の研究に従って，話題転換の機能を備えていない雑談対話システムを対象として，話題転換点の推定，転換話題の決定，転換文の生成という3つの部分から構成される話題転換機能を付加する．その中で，図 3.1 に示すように，本研究では，話題転換点の推定，転換話題の決定という2つの部分の処理方法を提案する．以下の提案手法では，Wikipedia^{*1}のデータから word2vec で学習した単語ベクトルを用いる．

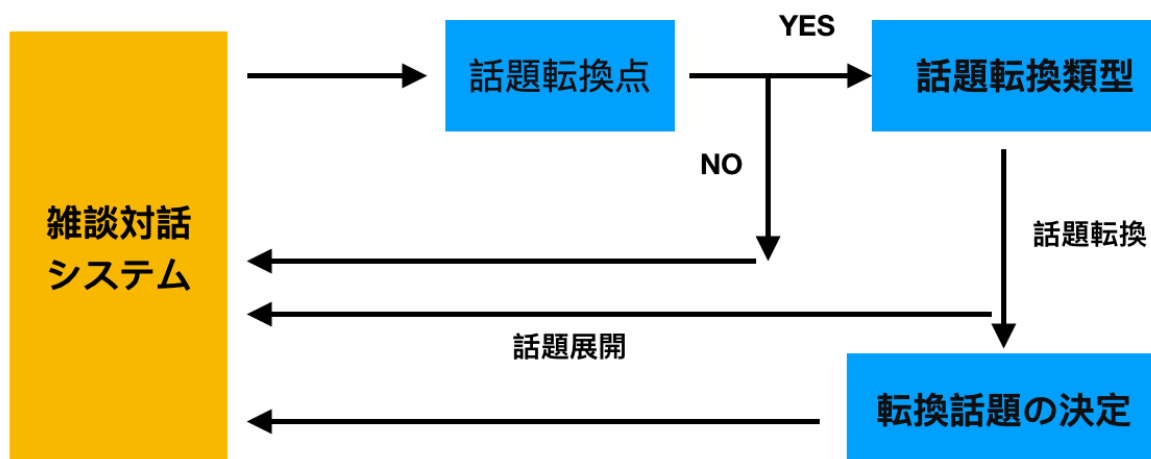


図 3.1 提案手法の流れ

3.1 発話文解析

発話文解析では，ユーザの発話を解析する．ユーザの発話を形態素解析器 Mecab^{*3}と辞書 mecab-ipadic-NEologd^{*4}で解析し，形態素を抽出と有効な発話文の選択を行う．

^{*1} <https://dumps.wikimedia.org/jawiki/latest/>

^{*2} https://ninjal-usamilab.info/lab/btsj_corpus/

^{*3} <https://taku910.github.io/mecab/>

^{*4} <https://github.com/neologd/mecab-ipadic-neologd/>

3.1.1 形態素の抽出

先行話題と弱関連する話題語を推定するため、発話文中の名詞を話題語とみなす。そのため、形態素解析器 MeCab と mecab-ipadic-NEologd を利用して、全ての「自立」、「一般」属性の名詞を話題語として抽出する。mecab-ipadic-NEologd は形態素解析エンジン MeCab に用いる単語分かち書き辞書で、週 2 回以上更新され、新語や固有表現に強く、語彙数が多いという特徴を持っているので、最大限に全ての話題語になれる名詞を抽出できる。また、「あと」、「うち」などの名詞は「一般」という形態素属性を持っているが、話題語として不適切なため、ストップワードとして除去する。

3.1.2 有効発話文の選択

雑談対話中で、相槌（例：「そうですね」、「いいですね」）や内容のない発話（例：「それは大変ですね」、「私もそう思います」）がよく出てくる。発話文間の類似度を計算するために、発話文を構成する内容語（名詞、動詞、形容詞）のベクトルが必要である。それを排除するためには、以下の条件を設定する。ユーザの発話文が、以下の条件 1 と 2 を同時に満たすか、もしくは条件 3 を満たすとき、有効発話文とする。

1. 発話文の文字数が 8 文字以上である。
2. 発話文に含まれる品詞の種類が 4 つ以上である。ただし、符号と感動詞は品詞から除く。
3. 発話文の名詞に対して、mecab の品詞細分類 1 が一般または自立の名詞である。ただし、ストップワードは除く。

3.2 話題転換点の推定

話題転換を行うためには、対話中の話題転換点の推定が必要である。先行研究 [8, 10] はシステムの発話文に対応したユーザの発話文だけを対象として、話題転換点を推定する。しかし、雑談対話として話題転換する前にそれ以前のユーザ発話も対象とすべきである。本研究では、ユーザの有効発話文の間の関連性を利用して、図 3.2 に示すようなユーザの発話文間 (u_1, u_2, u_3) の類似度を計算することで話題転換点を推定する。一方で、別所ら [11] の手法に従って、ユーザの有効発話文 (u_3) とシステムの有効発話文 (s_3) 間の類似度によって転換話題の類型を判断する。また、先行研究に関する話題転換の推定に用いて特徴と本研究で用いる特徴の違いを表 3.1 に示す。

表 3.1 話題転換点の推定に用いる特徴

研究	話題転換の推定に用いる特徴
堀内ら [8]	内容語数、一致度、連続性など
瀧口ら [10]	対話破綻の検出、応答時間
本研究	発話文間の類似度

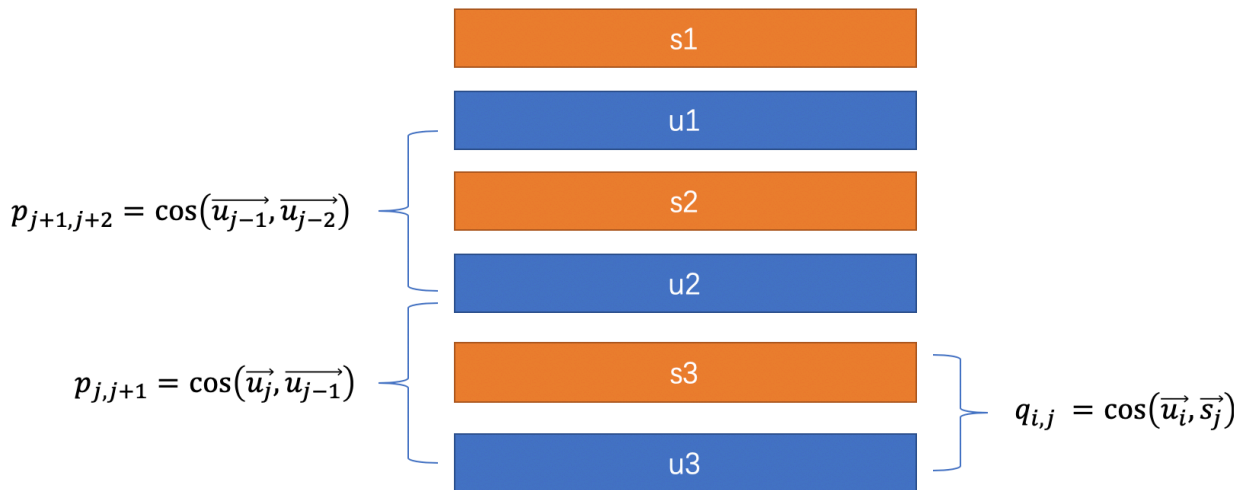


図 3.2 話題転換点の推定と転換話題の分類手法

3.2.1 話題転換点の推定

雑談対話中で、同じ話題について話す時に、ユーザの発話文間の類似度変換されると考えている。そのため、同じ話題に飽きる時に、発話中で話題語が少なくなる。それを指標として、有効発話文に基づき、話題転換点の存在を判定する。手順を以下に示す。

1. 発話文のベクトルとして、発話文中の全ての名詞、動詞、形容詞の平均ベクトルを計算する。
2. 直近の話題転換後からのユーザ発話列 u_T (T は直近の話題転換後の最初発話文) から u_i (i は現在の時点) に対して、隣接する発話同士の類似度 $p_{j,j+1} = \cos(\vec{u}_j, \vec{u}_{j+1}) (i > j \geq T)$ を計算する。
3. 隣接発話間の類似度 $p_{j,j+1}$ の平均値 p を求め、 $p \leq 0.22$ の時に話題転換を行う。それ以外は転換を行わない。

3.2.2 話題転換類型の推定

話題転換点を確定した上で、転換話題の類型を判断することが必要である。図 3.2 に示すように話題を続けるかどうかをユーザ (u_3) とシステムの発話文 (s_3) 間の類似度によって、話題展開 (中関連話題への転換) か話題転換を決定する。これに対応する素性として、過去 N 発話における話題転換点の 2 発話間の類似度を採用した。発話文関の類似度としては、発話文中で挙げた各素性ベクトルの種類ごとに、素性ベクトル間のコサイン類似度を取った。具体的な流れは以下で示す。

1. 直前の 2 発話 u_i, s_i に対して、隣接する発話間の類似度 $p_{i,j} = \cos(\vec{u}_i, \vec{s}_j)$ で計算する。
2. 隣接発話間の類似度 $p_{i,j}$ を用いて、以下のように判定する。
 - $p > 0.22$ のとき、話題を展開する。
 - $p \leq 0.22$ の時、話題を転換する。

3.3 転換話題の決定

転換話題の決定では、発話文解析で抽出された話題語に基づいて、転換前の話題語と弱く関連する話題語と転換話題とし、話題語候補の中で適切な転換話題を決定する。発話文中の話題転換類型は話題展開と話題転換を分けると、ここで話題転換類型だけに、弱関連話題の話題転換を推定する。以下で、2つの推定手法を提案することになりますが、両者から生成した弱関連話題の良し悪しを評価実験で検証すると考えている。また、先行研究 [8, 10] の手法を区別するためには、表 3.2 を示す。

表 3.2 転換話題の決定

研究	転換話題の決定手法
堀内ら [8]	ネットで人気話題のランダムに選択
瀧口ら [10]	検索エンジンの検索結果の特徴として選択する
本研究	話題語候補の類似度平均値と tri-garm の出現頻度

3.3.1 話題候補語の平均類似度による手法

有効発話文中の名詞を話題語として、各話題語との類似度が高い名詞を転換話題候補にする。そして、転換話題語候補それぞれに対して、全ての話題語名詞との類似度の平均値を計算し、最も低い平均類似度の話題語候補を、弱関連話題として転換話題とする。具体的な手順を以下に示す。

1. 直近有効発話文から名詞をすべて抽出する。例えば、表 3.3 に示すように、直近有効発話文の名詞「大学」、「食堂」、「銭湯」を抽出する。
2. 抽出した各名詞 u_i ($1 \leq i \leq n$) に対して、Wikipedia コーパス中の全ての単語との類似度を計算し、最も類似度が高い m 個の話題語候補 t_{ij} ($1 \leq j \leq m$) を求める。なお、本研究は $m = 30$ とする。表 3.3 に示すように、「学科, 大学院, 修士, 人文学科, 専攻」、「カフェテイナー, レストラン, 軽食, 調理場, 売店」、「サウナ, ホテル, 温泉, 混浴, 入浴」をもらう。
3. すべての話題語候補 t_{ij} に対して、 p_{ij} を次式で計算する。

$$p_{ij} = \frac{\sum_{k=1}^n \cos(\vec{u}_k, \vec{t}_{ij})}{n} \quad (3.1)$$

4. $p_{ij} > C$ を満たすもののの中で 最小の p_{ij} に対応する話題語候補 t_{ij} を弱関連話題とする。

表 3.3 話題分散

発話文の名詞	話題候補
大学	学科, 大学院, 修士, 人文学科, 専攻
食堂	カフェテイナー, レストラン, 軽食, 調理場, 売店
銭湯	サウナ, ホテル, 温泉, 混浴, 入浴

3.3.2 tri-gram の構造を用いた手法

発話文では、様々な出来ことや状態が述べられる。それらは主に「名詞＋助詞＋動詞」または「名詞＋助詞＋形容詞」という tri-gram で表れる構造を取る。そこで、類似した構造を持つ名詞を話題語候補として抽出する。その中にユーザに対話興味を喚起できる話題語があると考え。また、存在しない場合は、発話文中の名詞を利用して、先に名詞と tri-gram 確率が高いの「助詞＋動詞」または「助詞＋形容詞」を探す。それに基づいて続けて話題語候補を集める。収集した話題語候補と発話文の名詞との類似度を計算して、転換話題を推定する。具体の手順を図 3.3 以下

に示す．

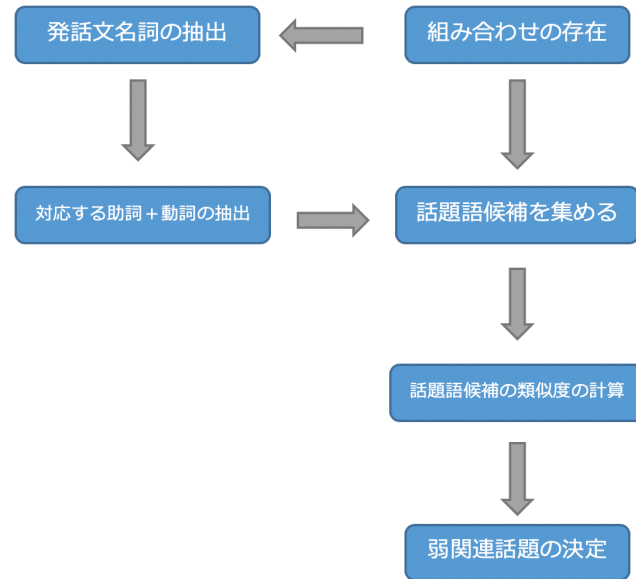


図 3.3 述語項構造を用いた手法

1. 発話文を tri-gram で区切り，「名詞 n_i + 助詞 q_i + 述語 v_i （動詞または形容詞）」という組み合わせを探す．そのような tri-gram 存在する場合は手順 3 に進む．
2. 手順 1 で条件を満たす tri-gram が存在しない場合には，2.1 節で抽出された各名詞 n_i に対して，tri-gram コーパスの中で，最も出現頻度が高い「名詞 n_i + 助詞 q_i + 述語 v_i （動詞または形容詞）」という並びの tri-gram を抽出する．例えば，表 3.4 に示すように，助詞「を」と動詞「観戦」または助詞「が」と形容詞「素晴らしい」の組み合わせがある．
3. 各名詞 n_i に対して，表 3.4 を示すように $\text{trigram}(w, q_i, v_i)$ の出現確率が最も高い m 個の名詞 w を話題候補語 t_{ij} とする．「を + 観戦」に対応する話題語「試合」，「野球」，「ゲーム」で，「が + 素晴らしい」に対応する話題語「天気」，「花」，「環境」である．
4. n_i と t_{ij} の類似度 p_{ij} を計算する．
5. $p_{ij} > C$ を満たすものの中で最小の p_{ij} に対応する t_{ij} が弱関連話題語とする．

3.4 話題転換文の生成

話題転換文の生成では，推定された転換話題を自然的に転換するために，適切な発話文を生成する必要がある．つまり，話題転換文とは先行話題から転換話題へ誘導する発話文である．本研

表 3.4 例：[助詞＋動詞] と [助詞＋形容詞]

推定した話題語	助詞	動詞
試合	を	観戦
野球		
ゲーム		
推定した話題語	助詞	形容詞
天気	が	素晴らしい
花		
環境		

究では，話題転換文の生成にテンプレートを利用する [8, 9]. 図 3.4 のような「相槌＋接続詞＋語尾」という生成パターンに基づいて，各要素をランダムに組み合わせて用いる.

相槌		接続詞		語尾
ええ	+	ところで	+	って知ってますか
そうですね		ではここで		に興味ある？
なるほど		さて		って説明できますか？
～え～		では		について話しましょうか？

図 3.4 話題転換文のテンプレート

4 評価実験

提案した手法を実装し，雑談対話データを利用して評価実験を行う．本研究では，話題転換点の決定と転換話題の推定の2つの処理に対して評価実験を行う．実験データには，人間どうしの対話履歴を収録した「BTSJ 日本語自然会話コーパス」をユーザとシステムの対話とみなして擬似的に用いる．推定された話題転換点と人手で判断した結果と比較して，適合率と再現率の値によって手法の有効性を検証する．また，人手で推定された転換話題と Twitter からランダムに選択する話題 [8] と比較して，推定した弱関連の話題転換の有効性を示す．

4.1 実験方法

話題転換点の評価実験（実験1）では，対話履歴データに対して，被験者2名両者が判断した転換点と提案手法で推定した転換点と比較して，再現率と適合率で転換点の性能を評価する．転換話題の評価実験（実験2）では，類似度の閾値 C を6つ設定する．実験1で人手で判断した話題転換点に対して，提案手法で決定した転換話題と生成した転換文を6名の被験者に評価してもらう．評価項目は話題の興味度，適切さと転換文の自然さであり，それぞれを1～5の5段階で評定してもらう．また，先行研究の手法 [3] と比較するために，Twitter のランキング話題語と一緒に評価を行う．また，話者の性別と関係の影響を排除するために，表 A.5 に示すように，話者の性別，話者間の関係，実験データを選択した．

表 4.1 評価実験員用いた対話データ

性別	関係	データ個数
男男同士	初見面	2
女女同士	初見面	2
男男同士	友達	2
女女同士	友達	2
男女同士	友達	2
男女同士	初見面	2

4.1.1 話題転換点の評価実験（実験1）

3.2 節で述べた手法で話題転換点を推定するためには，ユーザの有効発話間の類似度の閾値を計算すべきである．そのため，今回「BTSJ 日本語自然会話コーパス」から評価データ以外の対話履歴データ 20 件に対して人手で話題転換点を決定して，話題転換点までのユーザの有効発話文間の類似度の平均値を計算した．そして，話者の関係を考慮して，「初見面」と「友達」の各 10 個データに対して，有効発話文間の類似度の平均値を計算した．なお，以上に計算では，対話をしている二人のうちのどちらをシステム発話と見なすかで，2 種類の類似度を計算した．表 4.2 に示すように，「初見面」と「友達」の種類で分けて，各 10 個のデータによって話題転換点また話題展開点の有効発話文間の類似度を計算する．そして，「初見面」と「友達」のすべてのデータに対して，類似度の平均値は話題転換点の閾値とみなした．計算した結果は以下表 4.3 に示す．

次に，話題転点の適合率と再現率，被験者 6 名で対話履歴データ 12 件に対して，一人 3 件データをタグ付けてもらう．対話履歴のデータに対して，被験者 2 名両者が判断した転換点と提案手法で推定した転換点と比較して，再現率と適合率で転換点推定の性能を評価する．

表 4.2 データ閾値の計算結果

初見面	展開閾値	転換閾値	展開閾値	転換閾値
first-talk1	0.210	0.290	0.246	0.224
first-talk2	0.256	0.015	0.282	0.177
first-talk3	0.119	0.222	0.091	0.120
first-talk4	0.181	0.314	0.301	0.384
first-talk5	0.135	0.145	0.216	0.244
first-talk6	0.309	0.325	0.288	0.314
first-talk7	0.322	0.232	0.129	0.225
first-talk8	0.136	0.201	0.237	0.190
first-talk9	0.153	0.276	0.220	0.267
first-talk10	0.052	0.197	0.140	0.191
平均値	0.187	0.222	0.215	0.234
友達	展開閾値	転換閾値	展開閾値	転換閾値
friend-talk1	0.386	0.157	0.413	0.433
friend-talk2	0.245	0.259	0.323	0.176
friend-talk3	0.117	0.109	0.164	0.094
friend-talk4	0.505	0.101	0.262	0.176
friend-talk5	0.033	0.134	0.219	0.176
friend-talk6	0.363	0.227	0.001	0.183
friend-talk7	0.017	0.253	0.304	0.397
friend-talk8	0.186	0.140	0.201	0.286
friend-talk9	0.163	0.390	0.174	0.358
friend-talk10	0.209	0.198	0.145	0.095
平均値	0.222	0.197	0.221	0.237

表 4.3 データの種類による閾値の平均値

データ	展開閾値	転換閾値	(user,system を交替)	展開閾値	転換閾値
初見面	0.222	0.197	(user,system を交替)	0.221	0.237
友達	0.225	0.198	(user,system を交替)	0.228	0.224

4.1.2 転換話題の評価実験 (実験 2)

本研究の評価は、雑談対話システムを対象として転換話題の決定を提案したが、雑談対話システムが実装されていないため、人間同士の雑談対話データを利用する。評価実験では、話題転換点に対応する対話履歴を 6 名の被験者に読ませて、2 つ提案手法と先行研究の手法で生成された転換話題の興味度、適合率、転換文の自然さという 3 つの項目で評価してもらった。本研究手法の有効性を確認する。また、実験において類似度の閾値 C を 6 つ ($C = 0$, $C = 0.02$, $C = 0.04$, $C = 0.06$, $C = 0.08$, $C = 1$) に設定して、類似度の大きさが弱関連話題の興味度、適合率、転換話題が含まれる転換文の自然さにどんな影響を与えられるかを検討する。

4.2 実験結果

実験 1 の結果として被験者から人手で判断した話題転換点に対して、本研究の手法で推定した話題転換点の適合率と再現率を計算した実験結果は表 4.4 に示す。12 対話に対するマクロ平均的に計算すると、転換点の再現率は 0.054、適合率は 0.028 となった。結果が悪かったと分かった。

表 4.4 各データに対しての適合率と再現率

データ番号	被験者から転換点の個数	システム推定個数	再現率	適合率
1	5	7	0	0
2	5	11	0	0
3	7	3	14.2	33.3
4	6	6	16.7	16.7
5	2	3	0	0
6	9	12	0	0
7	7	15	0	0
8	6	17	16.7	5.8
9	4	14	25.0	7.1
10	2	15	0	0
11	8	12	12.5	8.3
12	14	26	7.6	3.8

実験2の結果を表4.5と表4.6に示す。これらの表では、本研究で提案した2つの手法について、6種類の閾値を用いて推定した弱関連話題の平均評価である。また、表4.7は提案手法と先行手法の比較結果（括弧内の数値は、6種類の閾値を用いて推定した話題語の中での最大値である）である。全ての評価項目について、提案手法の結果は先行研究の手法より悪い結果となったと分かった。

表 4.5 提案手法 (平均類似度) における評定平均値

	興味度	適切さ	自然さ
$C = 0.10$	1.85	1.63	1.83
$C = 0.08$	2.08	1.73	1.88
$C = 0.06$	1.84	1.47	1.85
$C = 0.04$	1.62	1.43	1.75
$C = 0.02$	1.80	1.45	1.67
$C = 0$	1.70	1.35	1.58

表 4.6 提案手法 (tri-gram) における評定平均値

	興味度	適切さ	自然さ
$C = 0.10$	1.84	1.65	1.93
$C = 0.08$	1.88	1.60	1.77
$C = 0.06$	1.61	1.56	1.70
$C = 0.04$	1.79	1.53	1.84
$C = 0.02$	1.71	1.55	1.73
$C = 0$	1.74	1.50	1.80

表 4.7 評価実験 2 における評定平均値

	興味度	適切さ	転換文自然さ
提案手法 (平均類似度)	1.81(1.85)	1.51(1.73)	1.76(1.88)
提案手法 (tri-gram)	1.76(1.88)	1.57(1.65)	1.80(1.93)
堀内ら [8]	2.24	1.69	2.05

5 考察

この章では、雑談対話システムにおける話題転換点の推定と転換話題の決定についての考察を行う。

5.1 話題転換点の評価

話題転換点には強関連話題への転換と弱関連話題への転換が含まれている。本研究は話題転換点のパターンを分析し、先行研究の手法と提案手法の違いを示す。

5.1.1 話題展開点の分類

話題転換点の特徴によって以下の9種類を分けている。

1. 挨拶完了の話題転換

雑談対話として、最初に必要な挨拶が欠かせない。そのため、自己紹介や挨拶などの社交辞令が終わる時点がよく話題転換点になる。以下の会話例で示すように、「はじめまして」から「よろしくお願いします」までは日常の挨拶と社交辞令である。それが終わると、次の話題に転換することがよくあると考えている。

A : はじめまして。
B : はじめまして。
A : えと、A と申します。
B : あ、A さん。
A : はい。
B : わたくし、B と申し (はい) ます。
A : よろしく申し上げます。
B : よろしく申し上げます。 (話題転換点)
A : 何をされていらっしゃるんでしょうか?。
B : あ、あの一、ふつうの会社員。

2. 話者間の相槌の類似度

雑談対話中で、話者間の相槌がよく出てくる。両者の相槌の一致度が高い場合は、話題転換点になりやすい。以下の会話例の中で、「そうでもないんですが」と「ないんですか」が大体一致している。

A: 何をされて (あ) いらっしゃるんでしょうか?。
B: あ、あの一、ふつうの会社員jというか
A: あ、会社員
B: 会社員で、新宿のほうで働いて (は一) おります。
B: はい。
A: 都心ですねー。
B: いや、《少し間》 そうでもないんですが。
A: ないんですか。 (話題転換点)
B: あ、お仕事されているんですか?。
A: はい。

3. 感動詞だけの発話文

感動詞だけで構成される発話文は話題転換点になりやすい。以下の会話例で示すように、対話文に感動詞「は一」や「へー」が出現する時には話題転換点を見なされやすい。

B: あ、お仕事されているんですか?。
A: はい。
B: は、何を?。
A: あ、教員やってます。
B: あ、こちらの大学で?。
A: はい。
B: あ、ほんとにですか。
A: はい。
B: あそうなんですか。
A: 見えませんか。
B: いいえ。

A: みんなからそう言われてますので。

B: はー。

B: へー。 (話題転換点)

A: 何の?、具体的には…。

B: あ、あの、食品会社で

4. 消極的な発話文の存在

雑談対話中で、ユーザが話題に飽きていることを示す表現として、消極的な発話文（例：「わかりません」、「苦手です」、「忘れちゃった」）が出てくる。以下の対話例で示すように、話者 B から「苦手です」という消極的な発話文が出現する時に、話者 A が話題転換文を発言している。

B: 渋谷とか行かれますか?。

A: ああ、渋谷はもう全然。

B: はー。

A: はい。

B: いや、私も、ああいう、ちょっと、若い (はい) 人がたむろするようなところは (うん)、あんまり、あんまり好きでないので

A: そうなんですか。

B: それと、同じような感覚で、(あー) 新宿も、ま新宿のほうはまだ、若干、大人の町というか (はい)、そういうところもあります、(ふーん) んですが、やっぱり、ごみごみしてるのはね

A: なるほど。

B: ちょっと、苦手ですね。

A: へー。

B: はい。 (話題転換点)

A: 今日どちらからいらしたん。

5. 自然な話題転換点

雑談対話中で、1つの話題が終了する時に、結論や事実や説明などの発話文が話題終了表

現として出現する。以下の対話例で示すように、話者 A が「雪が降ったときすごいですよね」という発話文に対して、話者 B が「ちょっと温度下がりますけどね」という結論または事実を話すことで、自然な話題転換が行われている。

- A: 夏は暑いし、冬は寒いっていう…。
- A: こっちのほうですか?、住んでる、あの,,
- B: ええ、こっちのほうです。
- A: 立川とか(ええ)って…。
- A: こちらでも、雪が降ったときすごいですよね、八王子とかって…。
- B: そう一、まあ、ここら辺まではね。
- B: もっとあっち行ったら、あの、ちょっと温度下がりますけどね。(話題転換点)
- A: 今日は、初めてそこの「路線名 1 線の誤り」線っていうやつ、「路線名 1」線ですか、あれ乗ったんですが、なんか途中ですごいい…。
- B: なんか「公園名」公園っていうのがありますね?。

6. 連続 3 回以上の相槌

相手の発話文に対して連続して 3 回以上相槌で返答する場合は話題転換点になる。以下の例で示すように、「へー」と「うーん、ふーん」、「いやあ」と「いやあ」、「はい」と「はい」という連続 3 回の相槌が発生する時には話題が転換される。

- B: あの、ガマとは関係ない?。
- A: あ、ガマとは関係ありますね、はい。
- B: はい、そうなんですか。
- A: ずいぶんお詳しいですね。
- B: 詳しくないです、ぜんぜん詳しくないです
- B: 詳しくりやもっと分かってます。
- A: あ、なるほどー。
- B: うーん、ふーん。
- A: へー。
- B: いやあ。

A: いやあ。
 B: はい。
 A: はー。 (話題転換点)
 B: え、今日も、あれですか?、藤沢のほうからです。
 A: そうです。

7. 発話文中の多様な相槌

「そうですか」、「なるほど」、「うん」などの相槌が多く含まれる発話文は、話題転換点になりやすい。以下の例で示すように、話者 A が感動詞「なるほど」と「へー」、相槌「そうですか」と一緒に発話することで、話題転換をしている。

B: 仕事で、あの、藤沢に住んでる方のところに (はい) 尋ねて行ったとき (はい)、なんとなく、気分としては日帰り出張っていうか。
 A: は、2 時間半ですから (ん)、お、し、ね、新幹線で大阪まで。
 B: そうですよ、そうですよ。
 A: ま、そう考えやそうかもしれません (ん) けど。
 B: でも、い、なんかの、海が近いっていう、ふうに思っただけで、なんか気分がよかったですけどね。
 A: しのまん中にいてもそんなに。
 B: そりゃそうでしょうけどね。
 A: なるほど、へー、そうですか。 (話題転換点)
 A: お仕事は何なんですか。
 B: 仕事ですか。

8. 外部環境や話者の表情と様子

雑談対話中で、今までの話題に関わらず突然の話題転換を行うことがよくある。それは話者間に共有されている言語以外の要素が影響するからであると考えられる。例えば、沈黙、発話待ち時間、話者外見、環境の変換などが挙げられる。以下の対話例で示すように、話者 A が「髪型変だ、シャンプーするとなー、や、お前にあの」という話者 B の外見について発話することで、話題を転換している。

B: ま、夏でしょ。

A: 夏だー。

B: ま、夏だね、これね。 (話題転換点)

A: 髪型変だ、シャンプーするとなー、や、お前にあの。 (話題転換点)

B: アイス食べてるよ、みんな。

A: ねー。

9. 内容が重複する連続発話文

雑談対話中で連続する発話文が同じ意味を持っている場合は話題転換点になる。以下の例で示すように、「たのしみー」と「わくわくするー」の意味が似ているため、話題転換点になると考えられる。

A: 《沈黙 2 秒》花火だそうだ、(うん) 花火買ってー。

B: たのしみー。

A: 楽しみだねー。

B: わくわくするー。 (話題転換点)

A: この天気が私の心を盛り上げない。

B: 何が?。

以上の話題転換点の分類と転換点の個数をまとめ、表 5.1 に示すように、各話題転換点種類の個数や話題関連性や転換文の話者を示した。9 種類の話題転換点に応じて、強関連話題または弱関連話題への転換が多いという傾向が見られる。その上で、話題転換点の発話文と新たな話題の転換文が同じ発話者の発話に含まれる場合（自己転換）と、両者の発話にまたがる場合（相手転換）がある。つまり、雑談対話システムの中で、システムがユーザと対話する時に、システムだけではなく、ユーザも話題転換文を発話する可能性があると言える。

また、話題転換点の種類と先行研究の話題転換点の推定手法で用いている指標の対応関係を表 5.2 に示す。堀内ら [8] の手法では、相槌の重複率（ユーザとシステムの発話文で一致する単語の割合）を用いて話題転換点を判定する。相槌に対して感動詞だけの発話文は発話内容が一番少ない状況だとみなして、話題転換点と判定する。3 回以上の相槌がある場合、連続 3 回の発話文が有効発話文ではない場合、話題転換点と判定する。外部環境や話者の表情や様子によると突然の転換

話題については瀧口ら [10] の対話破綻で対応することができる。

表 5.1 話題転換点の類型と個数

話題転換点の類型	弱関連話題への転換	強関連話題への転換	自己転換	相手転換	推定結果
挨拶完了の話題転換	0	1	1	0	
話者間の相槌の類似度	2	1	1	2	
感動詞だけの発話文	10	12	7	14	4
消極的な発話文の存在	5	3	5	3	
自然な話題転換点	12	6	4	14	2
連続 3 回以上の相槌	4	0	1	3	
発話文中の多様な相槌	7	0	3	4	
外部環境や話者の表情と様子	5	1	4	2	
内容が重複する連続発話文	2	0	0	2	
合計	47	24	26	44	

表 5.2 話題転換点の類型と転換指標

話題転換点の類型	転換指標	先行研究
話者間の相槌の類似度	一致度	堀内ら (09)
感動詞だけの発話文	内容数	堀内ら (09)
連続 3 回以上の相槌	連続性	堀内ら (09)
外部環境や話者の表情と様子	対話破綻	瀧口ら (16)
自然な話題転換点	発話文間の関連性	提案手法

5.1.2 話題展開点の考察

雑談対話データから分類された話題転換点が 9 種類存在する。本研究では、ユーザの連続発話文間の関連性が話題転換点の推定にどのような影響を与えるかを探求した。本研究で提案した話題転換点の推定手法で推定した転換点は表 5.1 で示すように、相槌が感動詞だけの発話文が 4 例、

自然な話題の終了が2例だけである。先行研究の手法では、自然な話題転換点の推定が難しいと考えられる。提案手法がこの類型で2例の話題転換点が推定できたことは、将来の雑談対話システムに対して、自然に結束する話題転換点の推定に役に立つ可能性があると考えられる。

表4.4で示したように話題転換点の評価結果は悪かったと分かった。具体的な原因を分析すると3つの原因が考えられる。1つ目は有効発話文の選択において、ストップワードとすべき単語を一般性名詞として有効発話文を判定することで、発話文間の類似度に悪影響を及ぼしている点である。2つ目は、有効発話文ではなくても話題転換点になれる。例えば、例文のように挨拶した後に話題転換する。

挨拶後の話題転換の例：

A：はじめまして。

B：はじめまして。

A：えと、Aと申します。

B：あ、Aさん。

A：はい。

B：わたくし、Bと申し(はい)ます。

A：よろしくお願いします。

B：よろしくお願いします。(話題転換点)

A：何をされていらっしゃるのでしょうか?。

B：あ、あの一、ふつうの会社員。

5.2 転換話題の考察

表4.7で示すように、提案手法で推定した転換話題の評価は、全体的に先行研究の転換話題よりも悪い結果となった。また、興味度と適切さに対して、提案手法の評価値が先行研究よりも高くなった転換話題の比率を表5.3に示す。興味度に関しては、先行研究よりも評価が高くなる転換話題の割合が低い。適切さについては、先行手法よりも提案手法のほうが良くなる転換話題が半数程度あることが分かった。

表 5.3 提案手法の評価値が先行研究よりも高くなった転換話題の比率

先行研究以上の項目	興味度平均値のみ	適切さ平均値のみ	両者
提案手法（類似度平均値）	0.26（8/31）	0.45（14/31）	0.16（5/31）
提案手法（tri-gram）	0.26（8/31）	0.52（16/31）	0.26（8/31）

以下の転換話題の例で示すように、雑談対話中の話題語に基づいて決定した転換話題がランダム話題より良くなった。その原因はネットの話題が時間や目標ユーザに対して制限があると考えている。雑談システムの中で、リアルタイムの話題だけで話題を転換することはユーザの生活に離れやすくなるので、ユーザの生活に関して一般的な転換話題も必要だと思われる。そのため、提案手法で推定した転換話題の適切さが先行研究よりも良くなったと言える。

転換話題の例:

発話文の話題語

‘言葉’, ‘言語’, ‘先生’, ‘学部’, ‘ディス’, ‘けい’, ‘金融’, ‘経済学部’, ‘教員’, ‘大学’, ‘国立’, ‘経済’, ‘プリン’, ‘学科’, ‘地元’, ‘場所’, ‘国際’, ‘フランス語’, ‘地域’, ‘専門’, ‘ドイツ語’, ‘最初’, ‘ことば’, ‘足下’, ‘分野’, ‘理論’

先行研究で決定した話題語

ジストニア

提案手法（類似度平均値）決定した話題語

法則市場, バンキング, ふるさと, ポテト, 平幕

提案手法（tri-gram）決定した話題語

足跡, 市街, 写真, 体調, 職務, 案件

そして、提案手法が先行研究の手法よりも悪い結果となった原因として、Twitter からの話題語が構造的に多様である（例「ローマの休み」、「終末時計」）からであることがあげられる。一方で、提案手法において、話題にはなり得ない単語（例：「せん」、「自体」）や単調な話題語がある（例：「一員」、「日替わり」）が話題語として出力されたことが評価を低くした原因だと考えられる。

5.2.1 話題展開点の類型によって転換話題の考察

表 5.4 に示すように、提案手法（類似度平均値）は「話者の相槌の重複率」と「突然な転換話題」という 2 つのタイプの話題転換点で、先行手法より興味度が高くなった。また、「感動詞だけの発話文」、「3 回以上の相槌がある場合」、「相槌言語が多い」、「突然な転換話題」、「連続発話文の重複」の話題転換点で、適切さに関して提案手法（類似度平均値）の方が先行研究の結果以上になった。

表 5.5 に示すように、提案手法（tri-gram）は「話者の相槌の重複率」と「突然な転換話題」という 2 つのタイプの話題転換点で、先行手法より興味度が高くなった。また、「話者間の相槌の一致度」、「3 回以上の相槌がある場合」、「相槌言語が多い」、「突然な転換話題」、「連続発話文の重複」という類型の話題転換点で、適切さは提案手法（類似度平均値）の方が先行研究の結果以上になった。

表 5.4 話題転換点の類型ごとの先行研究と提案手法（類似度平均値）の比較

話題転換点の類型	転換話題の個数	興味度（先行研究 / 提案手法）	適切さ
挨拶完了の話題転換	0	なし	なし
話者間の相槌の一致度	1	1.6/1.8	1.4/1.13
感動詞だけの発話文	10	2.22/2.00	1.58/1.58
消極的な発話文の存在	5	2.4/1.79	1.77/1.34
自然的に話題を終了する	8	2.37/1.82	1.76/1.41
3 回以上の相槌がある場合	1	2.17/1.64	1.17/1.36
相槌言語が多い	3	2.19/2.06	1.4/1.58
外部環境や話者の表情と様子	1	1.67/1.92	1.17/1.25
連続発話文の重複	2	2/1.83	1.3/1.3

表 5.5 話題転換点の類型ごとの先行研究と提案手法 (tri-gram) の比較

話題転換点の類型	転換話題の個数	興味度 (先行研究 / 提案手法)	適切さ
挨拶完了の話題転換	0	なし	なし
話者間の相槌の類似度	1	1.6/1.8	1.4/1.53
感動詞だけの発話文	10	2.22/1.91	1.58/1.47
消極的な発話文の存在	5	2.4/1.86	1.77/1.68
自然な話題転換点	8	2.37/1.78	1.76/1.41
連続 3 回以上の相槌	1	2.17/1.86	1.17/1.41
発話文中の多様な相槌	3	2.19/1.97	1.4/1.65
外部環境や話者の表情と様子	1	1.67/1.75	1.17/1.42
内容が重複する連続発話文	2	2/1.87	1.3/1.5

結果によると、「話者間の相槌の類似度」と「外部環境や話者の表情と様子」の 2 つ話題転換点のタイプは両方の手法について、興味度も適切さも先行研究よりも良くなったので、このような話題転換点における話題転換では、本研究の有効性を示したことになる。

5.2.2 不適切な話題語の排除

提案手法で決定した話題語の中に、話題にはなり得ない単語が含まれているために、興味度と適切さが下がっている可能性がある。そこで、以下に示す単語を人手で排除した上で、転換話題の興味度と適切さの平均評定値を求めた結果を表 5.6 と 5.7 を示す。これらの表から、話題にはなり得ない単語を排除することで、提案手法の興味度と適切さの評価値が上がることが分かった。

ワン、目前、直前、自体、せん、内的、前方、下方、下し、けい、同一、てい、きん、ところ、モノ、つけ、早め、ガタガタ、最小、いかん、たま、半々、誰か、キロ、不可、丁寧、切れ、もと、ふみ、あと

表 5.6 話題転換点の類型ごとの不適切な話題語の排除前後の提案手法（類似度平均値）の比較

話題転換点の類型	興味度 (排除前)	興味度 (排除後)	適切さ (排除前)	適切さ (排除後)
挨拶完了の話題転換	なし	なし	なし	なし
話者間の相槌の類似度	1.80	1.92	1.13	1.16
感動詞だけの発話文	2.00	2.03	1.58	1.61
消極的な発話文の存在	1.79	1.83	1.34	1.37
自然な話題転換点	1.82	1.84	1.41	1.42
連続 3 回以上の相槌	1.79	1.83	1.53	1.56
発話文中の多様な相槌	2.06	2.09	1.58	1.61
外部環境や話者の表情と様子	1.92	1.92	1.25	1.25
内容が重複する連続発話文	1.83	1.83	1.30	1.30

表 5.7 話題転換点の類型ごとの不適切な話題語の排除前後の提案手法（tri-gram）の比較

話題転換点の類型	興味度 (排除前)	興味度 (排除後)	適切さ (排除前)	適切さ (排除後)
挨拶完了の話題転換	なし	なし	なし	なし
話者間の相槌の類似度	1.80	1.92	1.53	1.64
感動詞だけの発話文	1.91	1.93	1.47	1.50
消極的な発話文の存在	1.86	1.93	1.68	1.71
自然な話題転換点	1.78	1.84	1.41	1.45
連続 3 回以上の相槌	1.82	1.85	1.39	1.39
発話文中の多様な相槌	1.97	1.97	1.65	1.67
外部環境や話者の表情と様子	1.75	1.83	1.42	1.50
内容が重複する連続発話文	1.87	1.86	1.50	1.50

5.3 今後の課題

本節では、話題転換点の推定や転換話題に決定に関する新しい手法を開発するための課題を検討する。本研究の手法における話題転換点を推定する時に、有効発話文の選択が重要である。実験結果を見ると、現在の選択方法がまだ不十分だと分かる。そのためには、発話行為の推定によって有効発話文を選択することが有効であると期待できる。南ら [12] はユーザの発話を分析し、対話行為にクラス分けすることは、ユーザの意図理解の 1 つとみなせると述べている。ユーザが挨拶をしているか、何を質問しているのかなどをシステムが理解することで、その後に話題転換するかどうかを決定できると言える。福岡ら [13] は発話の対話行為を推定する機械学習において、対話行為それぞれに対して有効な特徴を自動的に選択する手法を提案した。楊 [14] も話題転換の対話履歴を考慮して、話題終了表現の概念を提案している。

また、雑談対話システムにおける話題転換について、ユーザの興味度と適切さを上げるためには、発話文からの話題語抽出の性能向上と、多様な表現での転換話題の表現が重要だと考えている。丸田 [15] は Sidner の焦点モデルから対話の焦点モデルを提案し、深層学習を使って実装を行った。この技術を用いると、毎回発話文中の話題語が見つけやすくなると考えられる。転換話題の弱関連話題間で単純的な話題語を組み合わせると、多様な転換話題を生成させると考えている。例えば、本研究における話題語候補の中で「食堂」と「お替り」という話題語がある。これらの単一名詞を組み合わせると、「食堂のお替り」のような新しい話題語を生成できる。

6 おわりに

本研究では，ユーザの発話間の関連性を考えて，新たな話題転換点の推定する手法と，発話文中の話題語に基づいて弱関連話題への転換手法を提案した．評価実験 1 の結果より，発話間の類似度については，有効発話文の選択に応じる考えが不十分であることが分かった．表 4.7 より提案手法で生成した弱関連話題の興味度と適切さがないと分かったが，先行研究の手法で決定したランダム転換話題より悪くなると分かった．今後の課題として，ユーザ発話の発話行為を考えることで，有効発話文判定の性能向上が期待できる．また，推定した話題語を組み合わせることで，構造的に多様な話題語を推定すると，興味度と適切さを上げられる．

参考文献

- [1] 西村竜一, 原直, 川波弘道, 李晃伸, 鹿野清宏, 10 年間の長期運用を支えた音声情報案内システム「たけまるくん」の技術, 人工知能学会誌, Vo,28, No. 1, pp.52-59, (2013).
- [2] 太田 知宏, 鳥海 不二夫, 石井 健一郎, 発話生成を目的とした Wikipedia からの文抽出, JSAI2009, 2G1-NFC5-11 (2009).
- [3] Xiang Li, Lili Mou, Rui Yan, and Ming Zhang. StalemateBreaker: A proactive content-introducing approach to automatic human-computer conversation, *Proc.of Twenty-Fifth International Joint Conference on Artificial Intelligence*, pp.2845-2851 (2016).
- [4] 藤本 英輝, 高梨 克也, 河野 恭之, 木戸出 正継. 概念的関連性に基づく雑談の話題転換点分析, 人工知能学会全国大会, Vol,4, pp.188-188 (2004).
- [5] 中野 哲寛, 荒木 雅弘. 雑談対話システムにおける単語分散表現を用いた話題展開手法, 言語処理学会第 21 回年次大会発表論文集, pp.269-272 (2015).
- [6] 浅見一樹, 杉本 徹. 雑談対話システムにおいて話題転換に用いる連想語の取得に関する研究, 電子情報通信学会, 信学技報, Vol,117, No.82, NLC2017-11, pp.59-64 (2017).
- [7] 三上 佳孝, 萩原 将文. 対話システムにおけるランダム性を考慮した話題展開手法, 日本感性工学会論文誌, Vol,17, No.3, pp.365-373 (2018).
- [8] 堀内理沙, 上田祐大, 原田孝太, 韓東力. 雑談システムにおける話題転換, 言語処理学会第 18 回年次大会発表論文集, 2G3-01 (2009).
- [9] 前田和希, SongXin, 國政裕友樹, 豊田博之, 韓東力. 雑談システムにおける話題展開の性能向上, 言語処理学会第 16 回年次大会発表論文集, pp.250-253 (2010).
- [10] 瀧口 慈勇, 高間 康史. 対話エージェントにおける動向情報を用いた話題転換に関する予備的検討, 情報アクセスと可視化マイニング研究会第 14 回, Vol,14, No.2, pp.6-11 (2016).
- [11] 別所 克人, 東中, 竜一郎, 大塚 淳史, 牧野 俊朗, 松尾 義博. 雑談対話における話題継続願望判定の検討, 人工知能学会言語・音声理解と対話処理研究会, Vol,74, pp.1-6 (2015).
- [12] 南泰浩, 東中竜一郎, 堂坂浩二, 目黒豊美, 森啓, 前田英作. 対話行為タイプ列 Trigram による行動予測確率に基づく POMDP 対話制御. 電子情報通信学会論文誌. A, 基礎・境界, 95 (1), pp. 2-15 (2012).

-
- [13] 福岡 知隆, 白井 清昭. 対話行為に固有の特徴を考慮した自由対話システムにおける対話行為推定, 自然研究処理, Vol,24, pp.523-547 (2017).
- [14] 楊 虹. 中日母語場面の話題転換の比較-話題終了のプロセスに着目して, 国際交流基金日本語事業部世界日本語教育論集, No.17, pp.37-52 (2005).
- [15] 丸田要. 自然な対話継続のための推移する話題推定, 言語処理学会年次大会発表論文集, Vol,24, pp.528-531 (2018).

謝辞

本研究を行うにあたり，丁寧にご指導下さった内海彰先生，松吉俊先生，内海研究室の方々，評価実験にご協力くださった方々に深く御礼申し上げます．

付録 A 付録

A.1 ストップワード

あそこ, あたり, あちら, あっち, あと, あな, あなた, あれ, いくつ, いつ, いま, いや, いろいろ, うち, おおまか, おまえ, おれ, がい, かく, かたち, かやの, から, がら, きた, くせ, ここ, こっち, こと, ごと, こちら, ごっちゃ, これ, これら, ごろ, さまざま, さい, さん, しかた, しょう, すか, ずつ, すね, すべて, ぜんぶ, そう, そこ, そちら, そっち, そで, それ, それぞれ, それなり, たくさん, たち, たび, ため, だめ, ちゃ, ちゃん, てん, とおり, とき, どこ, どこか, ところ, どちら, どっか, どっち, どれ, なか, なかば, なに, など, なん, はじめ, はず, はるか, ひと, ひとつ, ふく, ぶり, べつ, へん, ペン, ほう, ほか, まさ, まし, まとも, まま, みたい, みつ, みなさん, みんな, もと, もの, もん, やつ, よう, よそ, わけ, わたし, ハイ, 上, 中, 下, 字, 年, 月, 日, 時, 分, 秒, 週, 火, 水, 木, 金, 土, 国, 都, 道, 府, 県, 市, 区, 町, 村, 各, 第, 方, 何, 的, 度, 文, 者, 性, 体, 人, 他, 今, 部, 課, 係, 外, 類, 達, 気, 室, 口, 誰, 用, 界, 会, 首, 男, 女, 別, 話, 私, 屋, 店, 家, 場, 等, 見, 際, 観, 段, 略, 例, 系, 論, 形, 間, 地, 員, 線, 点, 書, 品, 力, 法, 感, 作, 元, 手, 数, 彼, 彼女, 子, 内, 楽, 喜, 怒, 哀, 輪, 頃, 化, 境, 俺, 奴, 高, 校, 婦, 伸, 紀, 誌, レ, 行, 列, 事, 士, 台, 集, 様, 所, 歴, 器, 名, 情, 連, 毎, 式, 簿, 回, 匹, 個, 席, 束, 歳, 目, 通, 面, 円, 玉, 枚, 前, 後, 左, 右, 次, 先, 春, 夏, 秋, 冬, 一, 二, 三, 四, 五, 六, 七, 八, 九, 十, 百, 千, 万, 億, 兆, 下記, 上記, 時間, 今回, 前回, 場合, 一つ, 年生, 自分, ケ所, カ所, カ所, 箇所, ケ月, カ月, カ月, 箇所, 名前, 本当, 確か, 時点, 全部, 関係, 近く, 方法, 我々, 違い, 多く, 扱い, 新た, その後, 半ば, 結局, 様々, 以前, 以後, 以降, 未満, 以上, 以下, 幾つ, 毎日, 自体, 向こう, 何人, 手段, 同じ, 感じ

A.2 Twitter の話題

2020 年 1 月 24 日 12:30 までの Twitter のランキング

レヴナント, 離婚協議, ゴーカイブルー, サトヤスさん, 死者 25 人, 主婦休みの日, はぺこー, みのもんた, 終末時計, 東京ラブストーリー, EVO Japan2020, Wtrouble, USJ と京都, 表のあなたと裏のあなた, 業火の向日葵, バイオテロ, コナン映画, ローマの休み, ジストニア.

A.3 話題転換点の個数と対話履歴範囲

表付録 A.1 話題転換点の個数と対話履歴範囲

データ名	話題転換点個数	対話範囲（行数）
男男_初見面_01	3	1-55, 56-102, 110-176
男女_初見面_01	2	1-84, 117-310
男女_初見面_02	3	11-86, 87-123, 124-140
女女_初見面_01	7	1-56, 57-89, 90-135, 136-163, 164-292, 293-400, 401-435
女女_初見面_02	5	12-104, 105-192, 193-236, 237-518, 519-629
女女_友達_01	3	1-284, 285-711, 712-823
女女_友達_02	8	1-62, 88-133, 187-243, 287-340, 341-429, 470-519, 520-574, 575-614

A.4 類似度平均値で決定した弱関連話題

表付録 A.2 類似度平均値で決定した弱関連話題

男男_初見面_01						
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	左脳	お子様	本庁	パーフェクト	会意	ワン
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	実生活	大学院生	幕切れ	目前	少佐	直前
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	法則	市場	バンキング	ふるさと	ポテト	平幕
-	-	-	-	-	-	-
男女_初見面_01						
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	藻類	自体	朗報	蒼穹	タレント	文学部
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	テキスト	日替わり	片足	マガジン	売値	せん
-	-	-	-	-	-	-
男女_初見面_02						
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	ハイヤー	不動産	ルーム	三男	うめき声	顔ぶれ
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	口調	入り口	序ノ口	全社	自主	セーフティー
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	地面	商社	雪原	多面	場当たり	パーセント
-	-	-	-	-	-	-

女女_初見面_01						
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	単色	ビス	ファー	ゴセイ	主体性	内的
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	血液	園長	斜め	主将	前方	下方
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	自暴自棄	声楽	水中花	二等兵	ご破算	及び腰
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	サルーン	成鳥	尾羽	アメリカン	時点	下し
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	問柄	テーブル	信者	力感	せん	けい
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	エスカレーター	煮豆	全館	気位	打ち上げ	筋合い
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	神学校	高校	トレー	兵役	最小	ウェイ
-	-	-	-	-	-	-
女女_初見面_02						
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	若者	各校	気質	同一	ぜんまい	無法者
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	けむ	頭上	仰向け	茫漠	純然	惨憺
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	旧名	情緒	原液	合格者数	消石灰	硝酸カリウム
-	-	-	-	-	-	-

閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	長袖	貴女	教職員	産廃	間際	てい
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	学校	ページュ	すり身	レーシングカー	別個	ファイルサイズ
-	-	-	-	-	-	-
女女_友達_01						
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	まつり	ワッペン	古式	凶作	レッド	遜色
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	女優	プリマ	後手	持ち駒	後任	トレー
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	教場	本学	きん	歓声	穴熊	マウンテン
-	-	-	-	-	-	-
女女_友達_02						
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	尾びれ	忸怩	茫漠	絢爛	センター	顔ぶれ
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	実情	くだ	事案	動き	局面	情勢
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	軍帽	ガスマスク	終点	没年	域内	万物
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	てい	スト	歩幅	ダラー	行程	基線
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	天皇陛下	時差	陛下	フォーミュラ	茫漠	ベイ
-	-	-	-	-	-	-

閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	つま先	エディター	各人	バス停	地所	多数決
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	学生	イルミネーション	メロドラマ	純然	傘下	社説
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	紺色	古銭	血清	跡取り	御物	宝蔵
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	奥さん	けち	ブルー	ナイター	例年	川砂
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	文言	キャンピング	ラリー	セーフティー	サイドカー	サルーン
-	-	-	-	-	-	-

A.5 tri-gram で決定した弱関連話題

表付録 A.3 tri-gram で決定した弱関連話題

男男_初見面_01						
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	マナー	感想	ニュース	クイズ	主義	原因
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	ところ	励み	結婚式	手遅れ	ベテラン	上位
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	足跡	市街	写真	体調	職務	案件
-	-	-	-	-	-	-
男女_初見面_01						
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	開祖	地味	人形	等価	プレイヤー	一員
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	顔色	要因	例外	走り	裁判所	支払い
-	-	-	-	-	-	-
男女_初見面_02						
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	用事	個室	先客	仲間	感想	要素
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	要素	思い	モノ	プール	ハプニング	ページ
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	おなか	水槽	田舎	内部	対極	サイド
-	-	-	-	-	-	-

女女_初見面_01						
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	歌手	口座	キット	横道	サントラ	企業
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	ヒロイン	印象	商品	民族	宿命	都市
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	雲の上	箇所	性格	ワイン	プラン	製品
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	別れ話	つけ	片づけ	学生	買い方	積み立て
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	親戚	大げさ	アルバム	定額	早め	事務所
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	ガタガタ	ラスト	機会	土日	最小	都市
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	知識	レベル	文字数	いかん	情報	予算
-	-	-	-	-	-	-
女女_初見面_02						
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	作者	たま	地域	辺り	元旦	温泉
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	対象	性格	向き	ネット	事務所	早め
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	連中	半々	ノート	誰か	事件	資産

閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	想い	教室	感覚	ランダム	業務	買い取り
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	相手	キッチン	アニメ	棒状	キロ	匿名
-	-	-	-	-	-	-
女女_友達_01						
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	習慣	結婚式	走り	死語	目的	不可
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	盛り上がり	姿勢	傾向	兆し	戦い	証拠
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	喫茶	泳ぎ	カウンセリング	ディーラー	リゾート	バイキング
-	-	-	-	-	-	-
女女_友達_02						
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	形態	キャラ	大文字	フレンチ	丁寧	会社
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	踊り	コーナー	公園	もと	スピード	種類
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	ふみ	茫然	閑散	原料	職業	母語

閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	製品	内容	職業	レベル	見殺し	早め
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	口当たり	腹持ち	水疱瘡	中空	公園	惨事
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	ミルク	パスタ	ロープ	農薬	ガス	土日
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	誤字	やり方	側面	チケット	映画	余裕
-	-	-	-	-	-	-
閾値	$C = 0.1$	$C = 0.08$	$C = 0.06$	$C = 0.04$	$C = 0.02$	$C = 0$
転換話題	座敷	メンバー	寝たきり	名物	シリーズ	ワイン